国家社科基金资助期刊

图书情报知识

DOCUMENTATION, INFORMATION & KNOWLEDGE

主管／ 中华人民共和国教育部
主办／ 武汉大学

总第157期（双月刊）

全国中文核心期刊
CSSCI 来源期刊
中国图书馆学优秀期刊

2014



专业教育·我国LIS教育的现状及面临的挑战

我国LIS教育机构师资队伍建设的调查分析 肖希明 杨 蕾 4
我国LIS教育人才培养模式研究 唐 义 李金芮 11
我国LIS教育中实践教学的调查分析 肖希明 田 蓉 18
我国LIS专业毕业生就业情况的调查分析 李 硕 肖希明 26

博士论坛

科学数据权益分析的基本框架 顾立平 34
英美公共图书馆的社会角色设计 周 旖 52

图书、文献与交流

图书馆制度执行力概念、要素蠡测 白君礼 崔 旭 62
近五次我国全国国民阅读调查综述 卢 宏 71
图书馆阅读的行为角度研究——基于阅读困难群体问题的思考 王 虹 83

情报、信息与共享

基于访谈法的综合性文献数据库用户心智模型研究 韩正彪 90
图书荐购用户参与行为的调查与分析 李明鑫 田 丹 97

知识、学习与管理

政府门户网站公众满意度测评模型的实证研究 李海涛 102
基于案例分析的E-reserves版权问题研究 王 英 114
图书情报硕士专业学位 (MLIS) 教育发展状况调查报告(2013年) 段宇锋 董菲菲 121
全国图书情报专业学位研究生教育指导委员会2013年年会暨培养单位联席会议纪要 128

科学数据权益分析的基本框架

Basic Framework of the Analysis on Research Data Rights

顾立平
(中国科学院国家科学图书馆,北京,100190)

[摘要] 科研数据是科学发展的一个重要部分,而且作为一种知识资产,它也是图书馆知识管理的一项新兴挑战。科研数据权益是进行科研数据管理的关键,它的基本框架包括:利益相关方的权益、不同学科的管理方式、不同层级的管理政策、使用与引用的权益、存储与传播的权益、公共共享的权益等。在考虑现阶段我国科研数据管理与服务的需求后,最终提出面向科研工作流程的科研数据权益管理的未来工作。

[关键词] 数据密集型科学发现　开放科研数据　科研生命周期　数据生命周期　科研数据管理　知识产权

[中图分类号] G203　[文献标识码] A　[文章编号] 1003-2797(2014)01-0034-18　**DOI**:10.13366/j.dik.2014.01.034

[**Abstract**] Research data is an important part of science development, and as a kind of knowledge asset, it is also a new challenge on the library knowledge management. Research data is the key of research data management, and its basic framework includes the rights of stakeholders, the management approach at difference disciplines, the implementation policies at different levels, the rights of use and citation, the rights of deposit and communication, the rights of public sharing and so on. Based consideration of the present stage of needs of research data management and service, the article finally proposals a future work of the research-workflow oriented data right and interest management.

[**Key words**] Data-intensive scientific discovery　Open research data　Research life-cycle　Data life-cycle　Research data management　Scientific and technological information policy　Intellectual property rights

1　科研事业、科研数据与图书馆

发布科学理论及其实证和观察数据是其他人判断、同意、拒绝、理解该项工作的基础[1-2]。科研人员发布科研成果具有几个重要意义:①科学学说或者理论的成立,需要经过科学社群的认可,经过专家们对问题和解决方案的认识以及识别,得到认可;②科学是为社会服务、为人类全体福祉而不断自我追求和自我进步,因此科研成果应该尽可能地为公众所知并且利用在帮助人们的实践行动中;③就科学本身而言,构成科学的两大部分是机灵巧智的论述以及严谨可靠的证据,如同胡适所言:“大胆假设,小心求证”——思想和证据缺一不可。

英国皇家学会《科学是开放事业》中,针对现代科学发展所需的科研数据开放共享,写道:“数据必须容易接触而且便于找寻,让想检查数据的人能够清楚地理解数据,数据必须可以被评价,好让人们检查数据的可靠性和研究者的能力;数据也必须让其他人能够使用”[1-2]。之所以提出这条呼吁和要求,有其深远的科学文化断层隐忧,也有其对图书情报事业的期望与

[基金项目] 本文系中国科学院国家科学图书馆指向性人才研究基金资助项目“科技信息政策研究与咨询”(馆1203)的研究成果之一。

[作者简介] 顾立平,男,台湾台北人,博士,副研究员,研究方向:科技信息政策,Email:gulp@mail.las.ac.cn。

寄托。

科研事业的重要组成部分——学术期刊，源自早期十八世纪科学家们的通信。科研成果中的论述和证据两大部分（即：思想和证据）被同时刊载在后来逐步由皇家学会等机关所发展的这些科学通讯中。然而，随着科研发展的迅速扩展，许多证据（即：下文所述的科研数据）不得不被省略：比较一下孟德尔的遗传学说的证据和当前基因组序列的证据，就能明白纸本期刊难以刊载所有支持论文论点的数据。而且出于种种原因（如商业化期刊、科学评价、急功近利的科研人员等），使得一些科学不端行为混杂其中（如刻意不去呈现还存在若干不利论文论点的实验数据等）。

由此，导致了一个问题：构成科学理论的两大部分被一切为二，思想论述犹为可见，而实验数据则难见全貌。由此，又造成一系列的问题，例如：①独立存在的数据集如何证明普遍适用的科学理论命题；②除了少数的圈内人士，一般民众或者其它科研社群只能选择“相信”或者“不相信”某项科学论述，而无法亲自检验它；③假若选择相信，但是缺乏数据，人们并不能直接应用这项科研成果；④即使人们能够通过各种途径最终取得这些数据，许多数据并不能够被人们所解读；⑤即使能够解读，数据不一定是可以重复使用的格式；⑥即使人们取得可以解读并且重用的数据，不一定被授权使用；⑦在某些学科领域已经具有存储科研数据的系统平台，但是记载科学论述的论文和存储科研数据的地点并不一定相同，不一定提供方便可靠的链接；⑧潜在的跨学科研究所需的科学数据以不同格式、方式和规范存储在不同地点；⑨而且，在目前，缺乏科研数据所产出的问题是普遍性的，而提供科研数据促进学科发展的作法则是个别性的；⑩由于存在上述种种困难，使得新近科研人员容易误入歧途，忽视科研数据管理导致不当处理和不当呈现；⑪由于科研数据管理过于倾向信息系统化和既定的学科范式化，反而容易加深自我封闭的倾向；⑫科研贡献评价失衡：科学思想的影响被放大，而科学求证的贡献被低估，而恰恰两者是相辅相成的。

科研数据之于图书馆事业，并不只是“又一项要准备处理的信息资源”或者“又一种可以拿来用的情报服务或者知识服务”，而且还是图书馆事业支持科技创新和社会发展的当前以及未来重要推进方向。

2　我国的科研数据管理服务的实践综述

近年来，科研数据服务在我国图书馆界蓬勃发展，近期 CALIS 的几项研究成果，包括：对国外政策以及国内科研教育体系的科研数据服务发展[3]、汇整国外高校的科研数据生命周期的管理模型[4]、国内外高校数据共享平台的整理[5]、国内高校用户对科研数据管理的需求调查[6]、武汉大学图书馆实行科研数据管理的案例分析[7]，以及建立科研数据知识库的经验总结[8]等。这些内容是我国科研教育图书馆，发展与实施科研数据管理服务的重要基础。

科研数据权益管理是科研数据服务是否可持续发展重点。在科研数据政策方面，国内已有青秀玲老师翻译和引介 Christine L. Borgman 的研究论述[9-10]、著名信息搜寻行为研究者 Gary Marchionin 在武汉大学介绍 iSchool 的科研数据管理经验[11]、著名华人学者秦键在国家科学图书馆进行专题演讲[12]、黄永文老师等人介绍国家科学图书馆调研的开放科学数据政策、科学数据引用格式以及仓储案例等[13]，以及孟祥保和钱鹏老师等系统梳理国外数据教育的作法[14]等。这些成果是科研数据权益管理的重要基础。

据此，本文提出科研数据管理的两个主要面向：科研成果的数字数据长期保存（Preservation）和科研过程中的数字数据管理（Curation）。前者主要是将已经发表在学术期刊上的论文或者已经出版的专著中，所涉及的数据分析的数据集，予以描述、存储和提供下载与使用；后者主要是将个人、团体、机构或者受到资助的科研项目所产生的各种数据，在不同阶段中的各种处理的数据，从原始采集的观察数据，到汇总数据、检验数据、核实数据、剔除以及保留数据、转化为可以分析的数据集，以及最终经过整理的数据、简化后得以发表的数据等，均予以存储，并且提供得以后续分析利用的元数据说明、分析工具说明、分析结果说明等的一系列文件。采用第一种管理机制，需要考虑：①科研数据的层级和②科研数据的表述结构；而采用第二种管理机制，需要考虑①和②之外，还需要考虑到③科研数据的生命周期。

3　科研数据的基本概念

科研数据作为一种知识资产，既是“知识工业”的“原料”和“成品”，也是知识型社会中人们得以“证明”和“影响”他人对自身评价的一种资产，它的存储和传播方式已对知识型社会的组成产生影响。人类“社会”这个概念正在改变，从工业模型到知识模型的转变，最主要的因素是原料质变，也就是数据和信息交流方式的改变[15]。科技发展驱动社会变迁的其中一项影响即是传播方式的转变，不仅是在人群之中，也在科技研究的自身发展方式变化中。在《第四范式：数据密集型科研探索》这部网路开放的著作中，特别注意到科学自身如何在上个世纪末从实证主义型研究转向数据密集型研究：“科学传播，包括同行评审，也在发生根本性的变化。由于费用、时效、以及把实验数据和有关文件放在一起的需求，公共数字图书馆全面替换传统图书馆收藏出版物的角色”[16]。科研数据的层级[17]，如表 1 所示。在实证主义型研究的时代，只能从原始数据提炼为推导和组合数据，再形成科学文献内容，而在数据密集型研究时代，人们可以从科学文献链接到组合数据，并且从组合数据发现原始数据，进行数据之间的汇流与挖掘。

表 1　科研数据的层级

层级	定义
文献 Literature	从数据分析得到科学论点的数据
推导和组合数据 Derived and Recombined Data	经过处理能够被用来计算的数据
原始数据 Raw Data	从现实世界观察记录而来的数据

在数据密集型科研下，需要重新看待知识结构的关系，回归康德，界定科学[18]。如表 2 所示，事实或者数据是科研工作的根本基础，在事实以外或者缺乏数据佐证的情况下，人们无法进行科研成果验证的情况中，无法成立科学结果与论点。而对事实进行观察或者将数据进行挖掘的信息工作，是人们推进开放科学的主要工作，也是最后促成知识(知识，在此是指，人类行为准则和自然变化规律)产出的重要过程。因为科研数据的重要性透过新型科学研究技术而被放大，所以科研数据管理的工作，也成为一项重要的基础建设。

表 2　知识结构

知识结构	康德之问(Immanuel Kant)	经典著作
事实/数据	我能知道什么 Was kann Ich wissen?	《纯粹理性批判》(1781)
信息	我能期待什么 Was darf Ich hoffen?	《判断力批判》(1790,1793)
知识	我应该做什么 Was soll Ich tuen?	《实践理性批判》(1788)

当前正在快速推进的开放获取、开放知识、开放创新三种科技信息交流体系变化，为图书馆事业带来更多挑战和提出更高要求[19]。图书馆作为社会发展基础设施对新兴技术而言，既要作为公共知识供给的服务平台，又要成为各种科技创新投入施放的表演舞台。科研数据管理是图书馆发展的新兴机遇，图书馆对科研数据的发现、评价与遴选的能力，以及描述、组织、检索或者分析科研数据的能力，在未来至关重要。表 3 既是表述科学出版物的结构[20]，也是表述科研数据以建立数据元数据的重要内容[21]，它既是图书馆理论与实务中的 FRBR 和 RDF 等框架的设计基础，也是未来的科研数据的分析基础。表 1～表 3 是图书馆掌握科研数据的重要基础，此外，也是分析科研数据权益管理的轴线。

表 3　科研成果与科研数据的表述结构

	定义	例子
参照物	记载科研成果的实体	图书
标识符	指向参照物的符号	ISBN
元数据	描述标识符和参照物关系的内容	书名、作者、索书号

科研数据权益管理是科研数据服务是否可持续发展重点。短期来看，科研数据管理可以透过大型云

端设备或者小型数据库建立，但是长期来看，科研数据服务的营运与法规、技术、财务和组织[22]下的管理、需求、开放、成本、政策的支撑力度有关[23]。所以，科学数据权益管理框架的用途，在于能够按图索骥，在实践过程中，不断改善和深化相应的内容。

数据政策对科学家们以及它们如何持有科研数据等产生影响，而且来自资助机构和期刊的建议和强制执行会有更大的影响。数据政策要求接受资助者以及论文作者，确保作为出版前提的或者在项目中产生的数据可获取性[24]。科研数据具有三种存储、传播、引用和发布的形式，整理如表4所示。

表4　科研数据的三种存储、传播、引用和发布的途径[25-28]

三种情况	定义
科研数据独立发表	作为独立的信息对象，经由知识库而发布
科研数据与文本文献一同发表	此类情况又称数据论文
科研数据依附出版物发表	作为注释文本内容的材料，以丰富出版物内容的用途，作为一种说明文件得以发表（“使得出版物丰富化”）

科研方法和不同学科中的科研对象的特征，人们对科研数据会有不同的认识。科研数据的概念检验，必需服务不同学院和学术社群以及符合它们对科研数据概念的认识。同理，信息基础设施的需求同样来自于内容和用户需求。

作为科研过程的一部分或者结果的数字数据，科研数据出现在科研活动中的所有阶段里，从科研数据生成到科研结果产出，包括科学、社会科学实证研究或是文化现象观察等。为此，需要考察①利益关系人（Stakeholders）；②科研活动；以及③科研管理层级等，三项实际实施权益管理上的核心要素。

4　科研数据管理中的各方权益关系人

在科研数据权益管理政策中的具体措施，包括创建数据、数据处理、分析数据、保存数据、访问数据、重用数据等一系列科研数据管理工作流程中所需要的实施方案、记录规范、文件化方式、附加说明信息、传播政策声明与许可证制度等。

根据DCC的数据管理计划，实施流程包括六个阶段[29]。

（1）创建数据的实施方案：的设计研究、计划数据管理（格式、存储等）、共享计划、找到现有数据、收集数据（实验数据、观察数据、测量数据、模拟数据）、捕获和创建元数据。

（2）数据处理的记录内容：数据汇整（输入、数字化、转录、翻译）、数据清洗（检查、验证、清理数据）、匿名化数据、描述数据、管理和存储数据。

（3）分析数据的文件化过程：解释数据、导出数据、产生的研究成果、作者文章、准备保存的数据。

（4）保存数据的说明信息：数据迁移到最好的格式、数据迁移到合适的媒介、备份和数据存储、创建元数据和文件、数据归档。

（5）访问数据政策声明：分发数据、共享数据、控制访问、建立著作权与许可证、推广数据（主要应用在DigitalCuration的工作，在Digital Preservation则无需为数据扩散而实施执行工作）。

（6）重用数据的途径与许可范围：后续研究、新的研究、进行研究评论、审议结果、教育与学习。

然而，在上述的科研数据管理流程中，涉及至少六种类型的参与者和权益者，如表5所示。在科研数据管理的实施过程，需要尊重和保护这项具有不同的权利与利益的参与者和权益者。

其中，科研人员比较顾虑科研数据作为一种科学贡献的承认程度，以及引用科研数据的方式。学术社群通常关注科研数据的质量保证以及是否能够广泛为其它人所使用、重用和检验科学理论。科研机构一般着重科研数据作为一种机构知识资产的保存和管理机制，考虑开放共享来增加机构声誉。资助机构往往重视科研数据作为科研项目成果的一个部分，无论如何不能沦为以私人所有的方式挪作他用，或者宁可开放共享以增加资助机构在行业中的公信力。出版商经常关心以什么手段可以取得合理的商业利润，使得能够开发更好地科研数据服务。公众则关心是否有接近、使用科研数据的权利，以及知晓科研结果是否可靠的期望。

表 5　科研数据的参与者和权益者

利益相关方	科研数据的参与者
科研人员 (Researchers)[30]	科学家或者工程师(Scientists and Engineer) 研究生或者实习生(Students)
学术社群(Scholar Community)[30]	学协会(Learned Society) 专家团体(Expertise Group)
机构(Institution)[31]	高校(University) 科研院所(Institute)
研究资助者(Research Funder)[32]	国家及地方研究基金(State Funding) 独立法人团体(Independent Funding)
出版商(Publisher)[33]	商业出版商(Commercial Publisher) 非营利出版商(Non-Profit Publisher)
公众(Public)[33]	个人(Persons) 公众团体(Social Groups)

5　科研活动中的数据权益问题

社会资本(此处指:预期互惠)与科研社群对科研数据开放共享的行为方式有关,科研人员之间的竞争关系只是中等影响因素,真正决定科研人员对科研数据处理态度的关键因素是他们所在的社会网络与行业竞争体系[34]。不同学科的科研数据共享需求、文化和问题,各有不同,整理如表 6 所示。

表 6　不同学科的科研数据权益问题

学科	案例	权益问题
生命	可公开访问的 SNP 数据库(干细胞研究),允许个体识别的新方法和相关工具的应用[35]	以目前的技术保障措施,不能完全保护捐助者的身份资料,所以国际干细胞论坛(ISCF)发布“SNP 人类胚胎干细胞的出版政策声明”
声学	全球海洋观测[36]	部分地区噪音影响海洋生物的生存,正在建立数据采集、共享和建模的保障机制
化学	蓝方尖塔(Blue Obelisk)运动[37]	化学软件之间的互操作性、开放源码开发人员的合作,化学研究社群的资源和开放标准
医药	IMEX 联盟的的分子数据交换原则[38-40]	数据质量、化学生物学与药物发现的应用
基因	剑桥大学 CrystalEye 项目[41-42]	晶体结构数据的结构化 XML 的开放数据,提供浏览、搜索和新知快报
临床	心脏影像和临床数据的大型数据库(atlases)[43-44]	根据 Mozilla 公共许可协议版本 1.1 提供开源软件

不同学科的科研数据权益问题,环绕在两条主线上,一是在“最大程度地利用数据”和“最大程度地保护安全隐私”之间的平衡问题,这项挑战需要以发展电子信息管理的替代方案[36]在不同学科中进行解决;二是在“数据共享”和“数据获取”之间的权利义务关系[37]。前者涉及科研人员与一般人群之间,需要科研数据权益方面的基本权利保障机制,后者涉及科研人员与科研人员之间,需要科研数据权益方面的合作协调规范政策。

目前,上述这种争论已经日益获得各种共识,基本原则是科研数据的开放共享,但是在不同条件和情况下的具体实施作法,则是依赖在不同的科研层级之内的科研数据管理政策。这类政策包括重要但是经常为人们所忽视的权益界定和处理准则,这些涉及各方权益的政策,指引着利益关系人(Stakeholders)在保存、发布、使用和传播科研数据时,应该共同遵循的各项原则。

6　科研数据权益的管理层级

科研数据权益管理的层级,分为:国际、国家、机构、个人。如表 7 所示[1-2],不同层级关心以及能够进行的内容,有所差别。

表 7　科研数据权益管理的层级(参考[1-2])

层次	对象	案例
第 1 层	全球合作项目的国际案例	大型粒子对撞机、全球蛋白质数据银行
第 2 层	数据中心、国家机构	英国研究理事会、惠康基金会
第 3 层	个别机构、大学和研究机构	保存研究计划所产生的数据(彼此之间的差异很大)
第 4 层	研究小组或者个别研究人员	在传统的、现成工具的、小范围内里提交如 Excel 或 MATLAB 等数据

目前,在不同的权益管理层级中,出现各种出于不同目的而产生的数据政策。例如:欧洲议会和欧盟理事会启动的政府开放数据运动,主要根据 W3C 的数据描述建议,提供政府数据的访问方法,这个项目包括开放数据的结构与格式、可以被重用的数据,以及提供给公民和企业新服务[45]等,属于框架性质。英国地形测量局的 OS Open Data 在线地图门户网站,允许用户浏览、下载或者开发简单的数据应用,促进和鼓励更多地理信息数据注入[46],强化政府对支离破碎的地理位置(特别是边缘岛屿与海权领地等)的信息管治[47-48]等,属于国家政策。美国健康保险流通与责任法案(HIPAA)的隐私规则的要求,开放数据的管理技术必需采用模拟攻击和匹配试验的风险识别,经过鉴定,能够实现开放数据的建立原则[49]后才能实现,属于行业内的技术细节。而化学信息学工具,对分析化学、生物化学药理学和药物发现等的应用[50],则属于研究团队的项目工作。

最近,国际间的科研管理政策,具有一种朝向科研数据开放共享的趋势,例如:欧盟成员国有政策实施"公共资金资助科研成果的科研数据通过电子基础设施实现公共获取、使用和重用"[51]。美国白宫科技政策办公室 (OSTP)的行政命令是:"无论全部或者部分受到联邦资助的科研项目,所产生的数字形式的科研数据应该进行存储,并且提供得以搜索、检索和使用的公开获取[52]"。欧盟计划在第八次 HORIZON 2020 框架中加入类似要求[53]。这些政策在很大程度上确立了国际间将朝向科研数据开放共享的趋势,然而,具体如何进行,则涉及到一系列诸如:使用权益与引用问题、存储权益和采集问题、传播权益与发布问题,以及共享权益与政策问题等。

7　科研数据权益管理的使用权益与引用问题

数据资源的可用性(usability)包括:反应具体问题、成为可理解可操作的数据、具有足够的透明度和清晰度,以及支持用户互动与访问[54]。科研数据权益管理必需在技术、财务、法规和组织能力上,解决这些问题。

首先,在四十年前发表一篇论文需要包括完整的数据以提供重复操作,然而鲜有期刊可以这么做到,因为新的科技研究所需的大量数据远非期刊能够刊载。这使得科学成果的两个主要部分:思路和证据,被分开来,形成了一道数据鸿沟(data-gap),不利于科学的自我审查机制。试问:如果不能访问和评论科研数据,那么如何挑战和修正科学理论呢?

要解决这个问题,并且能够产生在数据密集型科研中的科学加速进步的效益,需要科研数据的开放共享。表 8 显示了科研数据开放共享的程度。只有在可访问性、可评估性、可解读性、可使用性等四个条件都具备的情况下,科研数据能够被重复验证,以及据此形成科学探索与发现的新知识。

表 8　科研数据开放共享的程度(参考[1-2])

层级	定义
accessible 可访问性	数据能够在某个地方被找到。(可读性)
accessable 可评估性	数据能够被解读而且被检查。(可靠性)
intelligible 可解读性	数据能够被理解、解释和判断。(完整性)
useable 可使用性	数据或者信息能够被任何人使用的形式。(重用性)

若将科学论文与支持论文结果的实际数据和元数据紧密集成,则意味在论文、数据、元数据(描述论文与数据)以及链接上,进行有效而且长期维护的能力。这类科研数据的特殊要求,有别与传统期刊元数据和文摘的工作流程。目前,国际晶体学联盟(International Union of Crystallography,IUCr)的学术期刊,已

经进行若干尝试，它将科学论文与实验数据的元数据集成[55]，提供短篇结构文章、实际数据文档、含有讨论的文本批注、传输中介以及工具等，从而可以便利地创作和创建交互式分子图形等服务。

另一方面，如果要求科研人员提供数据，那么必需保障他的权益：保障科研人员提供科研数据也作为一种科研产出的贡献。如同期刊论文无论是放在何种平台，均能按照刊名、卷期、页数、署名等被引用或者证明其工作成果一样，科研数据也需要能够被引用和被识别的机制。

科研数据的引用方式有两种取向，一种是注册单一数据集，即赋予每个数据集一个可以标示符号，作为引用来源；另一种是将数据集作为论文的一种附加材料。

德国科研基金会 DFG 的"科研数据的发布与引用"计划曾经以 DOI 域名和 URNs 注册的基础设施来处理科学数据，作为识别符的 DOI 域名应用，可以链接一篇科学论文和文章中所分析的科研数据[56]。文章和数据集通过他们各自的 DOI 域名，能够共同被引用。这种引用和链接的形式提供科研数据，直接指向科学论文，也就是 supplementary data（附加数据）。范例如下[57]：

Kuhlmann，H. et al.（2009）：Age models，iron intensity，magnetic susceptibility records and dry bulk density of sediment cores from around the Canary Islands.

PAMGEA，Bremen

doi：10.1594/PANGAEA.727522

指向如下论文：

Kuhlmann，Holger；Freudenthal，Tim；Helmke，Peer；Meggers，Helge（2004）：Reconstruction of paleoceanography off NW Africa during the last 40,000 years：Influence of local and regional factors on sediment accumulation. In：Marine Geology，207（1-4），209-224

doi：10.1016/j.margeo.2004.03.017

在此基础上，众多科技图书馆和信息中心在 2009 年 12 月共同成立 DataCite 为一项目标而努力：让科学家在互联网上访问科研数据、接受科研数据作为独立而且受到引用的科研对象，以及确保良好的科学实践规则[58]。它的三项主要工作，包括[59]：

（1）作为数字资源唯一标识符（DOI）的官方注册机构。

（2）积极发展标准规范、最佳实践交流，以及制定工作流程。

（3）建立可供访问的中央元数据库，实施开放共享。人们可以通过第三方（如 WOS）进入这个数据库，免费使用其中的内容。

这种方式，在基本权利上，保障科研数据在科学交流体系中的地位。科研人员发布科研数据的贡献，得到承认。首先，不同程度的科研产出的科研数据（参考表 1），作为参照物，具有标识符和元数据（参考表 3），能够被承认和引用。其次，各种学科内的科研数据（参考表 6）能够据此形成各自的工作流程，以及逐步形成行业规范。第三，各种数据银行的内容，可以被检索和使用。最后，不同层级的科研数据（参考表 7）均能纳入这个权益管理制度内。

8 科研数据的存储权益和采集问题

数据管理工作流程[60-61]、科研工作流程[62-63]、科研数据管理工作流程[64-65]、科研数据权益管理流程，如表 9 所示，具有不同内涵。

据此，得以分析科研数据管理的两个主要面向：科研成果的数字数据长期保存（Preservation）和科研过程中的数字数据管理（Curation）。前者主要是将已经发表在学术期刊上的论文或者已经出版的专著中，所涉及的数据分析的数据集，予以描述、存储和提供下载与使用；后者主要是将个人、团体、机构或者受到资助的科研项目所产生的各种数据，在不同阶段中的各种处理的数据，从原始采集的观察数据，到汇总数据、检验数据、核实数据、剔除以及保留数据、转化为可以分析的数据集，以及最终经过整理的数据、简化后得以发表的数据等，均予以存储，并且提供得以后续分析利用的元数据说明、分析工具说明、分析结果说明等的一系列文件。

表 9　科研数据权益政策的概念与操作化定义

关键概念	本文定义
科研数据流程	根据信息生命周期，进行采集、标示、存储、保存、释出/使用、剔除等
科研工作流程	根据科研工作本身的周期性工作要点，进行记录与管理
科研数据管理工作流程	根据科研工作中，每个阶段所产生的数据，进行各种不同方式的管理
科研数据权益管理流程	在科研数据流程与科研工作流程彼此交替发展的过程中，规范与协调各方参与者与利益相关方的潜在冲突，使得科研数据管理工作流程的记录、存储、保存、发布文件的措施得以顺利运行

参考数据档案管理的生命周期[66]，以及表 7～9 的内容，制定表 10 科研数据权益管理的流程，分为三个维度：科学数据流、利益相关人物，以及权益管理。

表 10　科研数据权益管理的流程

维度	目标	任务
A. 科学数据流	A-1 创建数据	A-1-1 设计研究 A-1-2 计划数据管理（格式、存储等） A-1-3 共享计划 A-1-3 找到现有数据 A-1-4 收集数据（A-1-4-1 实验、A-1-4-2 观察、A-1-4-3 测量、A-1-4-4 模拟） A-1-5 捕获和创建元数据
	A-2 数据处理	A-2-1 输入数据、数字化、转录、翻译 A-2-2 检查、验证、清理数据 A-2-3 匿名化数据 A-2-4 描述数据 A-2-5 管理和存储数据
	A-3 分析数据	A-3-1 解释数据 A-3-3 导出数据 A-3-4 产生的研究成果 A-3-5 作者文章 A-3-6 准备保存的数据
	A-4 保存数据	A-4-1 数据迁移到最好的格式 A-4-2 数据迁移到合适的媒介 A-4-3 备份和数据存储 A-4-4 创建元数据和文件 A-4-5 数据归档
	A-5 访问数据	A-5-1 分发数据 A-5-2 共享数据 A-5-3 控制访问 A-5-4 建立版权 A-5-5 推销数据
	A-6 重复使用的数据	A-6-1 后续研究 A-6-2 新的研究 A-6-3 进行研究评论 A-6-4 审议结果 A-6-5 教育与学习
B. 利益关系人	B-1 研究人员	B-1-1 科学家或者工程师 B-1-2 研究生或者实习生
	B-2 机构	B-2-1 高校 B-2-2 科研院所
	B-3 学会和专业团体	B-3-1 学会 B-3-2 学术和专业团体
	B-4 公共研究资助者	B-4-1 国家及地方研究基金 B-4-2 独立法人团体（国家入股 50% 以上）
	B-5 科学期刊出版商	B-5-1 商业出版商 B-5-2 非营利出版商

续表 10

维度	目标	任务
B. 利益关系人	B-6 企业研究资助者	B-6-1 企业投入研究 B-6-2 企业委外研究
	B-7 政府	B-7-1 产业部门 B-7-2 监督部门
	B-8 公众	B-8-1 个人 B-8-2 公众团体
	B-9 数据生产者	B-9-1 数据采集者(Data originators) B-9-2 数据提供者(Data providers)
C. 权益管理	C-1 公共知情权利 C-2 合法商业利益 C-3 国家安全机密 C-4 个人隐私权 C-5 知情同意权	
	C-6 数据开放性	C-6-1 可访问性(accessible) C-6-2 可评价性(accessable) C-6-3 可解读性(intlligible) C-6-4 可使用性(useable) C-6-5 可操作性(executable)
	C-7 数据保密性	C-7-1 匿名化(Anonymisation) C-7-2 有限访问(limited Access) C-7-3 封闭化(Closed)

9 科研数据的传播权益与发布问题

若将科学论文与支持论文结果的实际数据和元数据紧密集成，则意味在论文、数据、元数据(描述论文与数据)以及链接上，进行有效而且长期维护的能力。如表 11 所示，包括一系列权益管理政策，未来可预见的科研数据发布系统，宜先进行政策制定工作，方能规划系统平台的存储要求、权限管理和传播许可制度等。

表 11 科研数据发布系统的数据权益管理

科研数据发布系统	科研数据服务的功能	涉及的权益管理政策
开放数据[67]	提供学科领域内的数据挖掘与知识发现	隐私权规范、安全性准则
开放系统架构[68]	提供不同设备、平台、软件的数据交互	数字著作权管理(DRM)
开放关联数据[68]	跨越不同的数据源，获取所需数据	标准协议、获取、许可证
开放平台[68]	提供数据接口，对接应用程式界面	数据库权(DR)与创作共用
知识元库[68]	论文与论文、论文与数据、数据与数据	许可条款的许可证编码
数据价值计量[69]	科研论文与数据的全谱段使用统计计量	终端用户的使用记录和保密

在科研数据发布与链接服务的方向上，有两大重要技术线路：关联开放数据和知识元库。尽管技术已经成熟，但是决定是否能够实现的关键因素是数据权益管理。

(1)关联开放数据与传统意义上的搜索不同，它能够跨越不同的数据源，在诸如 RDF 格式和 SPARQL 查询语言标准的基础上，执行复杂查询，它更类似于分布式或者联合数据库，但与之合作的数据源则是独立维护和更新他们各自的数据[70]。将开放信息转换为可计算的开放知识，有关工作包括[71]：①支持开放化信息的可重用；②支持可重用信息的结构化；③支持结构化信息的语义化；④支持语义化信息的关联化。关联开放数据(linked open data，LOD)过程包括：采集数据、公布关联数据、规范化链接中的“连接点”、从个别数据集获得数据，以及客户端应用程序数据等[72]一系列处理程序和规则。然而，开放关联数据，

包括两个部分[73]:开放(标准、获取、许可)和关联(识别符、数据模型、本体、查询语言)。

(2)知识元库(Knowledge base, KB):受到知识元库开发者所编译、分发、维护的庞大数据库,其有关电子资源的信息,例如标题列表、覆盖日期、无缝链接等;早期是为“动态更新链接[74-76]”与“执行多方链接[77]”而设计,后为解决“适当副本(Appropriate copy)”问题[78]。根据同样原理,可以开发学科内用于数据密集型科学的知识元库,如“生物标识语言(Biological Expression Language)”[79],以及,智慧链接各种科研数据中心的数据库,提供检索、访问、下载与计算等功能,如表 12 所示的 Altmetric 等[80-84]计算使用数据的方式。目前,知识元库的供应链是由一系列复杂的角色所组成:出版商、其他内容持有者、订阅代理者、链接解析器供应商、图书馆等[85-86]。美国 NISO 正在促进新的标准,解决数据质量和许可证的问题[85]。这些问题也会发生在科研数据在学术交流体系中的传播制度。

表 12　在 Altmetrics 的五种 ID 参数[80-84]

ID 标准	常见的数据中心	在 Altmetric 的范例
Altmetric article ID	Altmetric	http://api.altmetric.com/v1/id/241939
DOI.	DataCite	http://api.altmetric.com/v1/doi/10.1038/news.2011.490
PubMed ID	PubMed	http://api.altmetric.com/v1/pmid/21148220
arXiv ID.	arXiv.org	http://api.altmetric.com/v1/arxiv/1108.2455
ADS bibcode	NASA Astrophysics Data System	http://api.altmetric.com/v1/ads/2012apphl.100y3104b

如果寻找专业技术供应商,能够迅速搭建平台,然而,根据国外研究显示,一般(其它行业的)专业技术供应商如果愿意提供开放平台,捍卫其知识产权的方式会更加猛烈[87]。这是因为目前广泛采用的软件密集型系统是一个开放架构(OA)加上多个开放 API 作为开放源码软件(OSS)的部件或组件;其优点是实现开放的众多好处,然而其挑战是,各个组件都受到不同的著作权或者知识产权许可证的约束[88]。故而,无论是否建立平台,建议提前准备科研数据权益管理政策的内容制定。表 13 显示科研数据发布涉及的权益关键问题。

表 13　科研数据发布的重要权益概念

术语	定义	争议
技术保护措施(TPM)[89-90]	通过硬件或软件或者综合两者,以防止或限制复制,以及控制访问和数字内容的技术手段	以技术手段限制原本不应或者不被允许的限制范围,从而获得商业利益
数字著作权管理(DRM)[91-92]	在格式转换和档案迁移上,透过多重签名认证手段,达到对用户使用内容的限制	与开放平台或者开放数据的理念相违背
数据库权(DR)[93-94]	授予一个数据库(通常是出版商)拥有者在大量投资后,获得文稿内容的独占权(类似复制权的制造商著作权)、播放权(如向公众传播权),以及发行权等	“下载”是服务(免费服务、开放服务、收费服务、加值服务)或者是权利(行使权利、有偿权利、无偿权利、公共权利、限制性权利等)尚未得到充分共识
创作共用(CC)[95-96]	解决传统著作权模式在互联网上所面临的困难(例如,允许非商业性质的复制和传播等)。不过,创作共用协议是不可撤销的,也就是创造者不能阻止已获得 CCL 的副本流通	在网上作品受著作权法规定的法律环境(例如,实际上无法立法明确商业目的和复制行为的定义),与社会规范和期望(例如,用户在工作环境中的自由活动等)时有冲突
科学创作共用(SCC)[97]	旨在消除在科技创新合作中,不必要的法律和技术壁垒。其长远目标是简化研究过程中出版、许可和获取	用于研究人员之间的数据共享,在信息技术上、合作方式上、运行机制上还有许多困难

10　科研数据的共享权益与政策问题

科研数据管理,并不仅仅是分布式网路系统和个性化权限管理的问题。与过去的分布式异构数据的不同之处在于,需要更简单、更灵活、更有效的技术,以及,更重要的是,有一个技术和文化适应的需求[98]。在生命科学领域中,关联开放数据(LOD)的技术问题

并不足以影响建设以数据为中心的“大生物学”学科，然而，在异构数据的文化上缺乏了解[99]：①遵循一种数据标准；②科学家共享数据；③激励措施等的问题，使得相关工作的进展存在无形障碍。为了使研究数据的充分利用，生物科学社群开始倡议技术和奖励机制，以支持互操作性，促进开放科学与文化的增长等。

强化元数据标准的发展，可以提高数据的访问，但是更重要的是数据权益管理的部分，例如工作流程、数据重现性、所捕获的数据源、数据保存和复原、使用的归属和确认等一系列问题[100]。以结构生物学为例，早期发展阶段并没有维护实验数据的制度，然而，随着科研问题日益复杂化，目前在结构生物学界的人们越来越相信这些未被充分保管好的数据，对于每个大分子结构的研究都具有重要的参考价值[101]；数据被存档和收集起来，并且在适当时机发布，是开放科学的重要举措，这涉及到科研数据的保存、流通、管理、保障等一系列的政策问题。

另一方面“开放科学（Open science）”在科学界兴起，意味：除了最终研究成品之外，研究人员分享他们在研究过程中的每个元素，以促进研究人员之间合作的一种新型科研方式，它能够促使科研人员自发性创建新的虚拟合作研究[102]。然而，在高成本的研究工具以及仅有少量研究成果能够获利的情况下，在生物医学研究领域内的知识产权（IP）制度不断扩张到学术圈内[103]，知识商品化现象显著[104]。研究人员目前处于一个混合“开放科学”和“专利制度”的体制之内。科学界往往对于开放科学有着不同的认识，例如什么是重复性、计算资源、科学荣誉、数据共享、成果共享、同行评审、成效评估等的含义往往模糊不清，从而需要广泛的意见和动机调查，并且通过不同的宣传，达到凝聚共识和达成协议[105]。这些工作有赖专业科技信息政策研究与咨询的工作小组推动，方能使得科研人员更容易获得学术信息交流的便利服务，以及最好地从中受益。科研数据权益与政策见表14。

表14 科研数据权益政策工作[1-2]

1	科学家传播所收集和建模的数据并且使其自由和开放获取
2	大学和科研机构在支持开放数据的文化中扮演重要角色
3	评估大学研究，注重对出版物的开放数据进展，建立评价
4	学会、科研和专业团体促使成员成为开放科学的先驱者
5	研究理事会和慈善机构资助的项目应提升科研数据的交流
6	科学期刊加强要求支持文章论点信息的可访问、可评估、可使用和可追溯
7	工业和监管部门共同确保公众利益的数据、信息、知识共享
8	政府重视开放数据和开放科学提升卓越科学基础的潜力
9	数据集在一种按比例治理的体系内管理；具有高度公共价值的研究潜力的情况下，个人数据才能被共享
10	基于现有商业标准调整更加宽泛的信息共享协议和实践
11	开放性在法律保护商业价值、隐私、安全和保密的范围内

11 我国实施科研数据权益管理的需求

数字科研数据因为学科和方法而有不同的数据类型、不同程度的集成以及数据格式。就科研数据使用与重用的获取目的而言，假若缺乏元数据以及描述内容和工具如何创造、存储、调整和分析的合适文件，则数字科研数据丝毫没有价值[106]。

目前，我国已经开展许多科研数据管理、科研数据管理服务，以及科研数据管理服务教育的措施。鉴往知来，从中国科研数据管理的三个阶段（表15）以及中国图书馆界的发展阶段（表16）的梗概，可以得知实施科研数据管理时，有其权益政策的服务需求。

目前，中国科研数据管理位于第二阶段，科研数据仓储的建设已有相当基础，并且在数据质量控制上也有若干规范，正在朝向科研数据开放共享的权益平衡方向发展中。而图书馆在科研数据管理的服务开发方面，正在处于第一阶段和第二阶段的过渡时期，由于已经认识到科研团队对于图书馆协助他们进行数据管理的服务需求，以及从图书馆发展战略的角度上，认识到图书馆作为科研数据仓储建设和质量控制等机制的必要性。

表 15 中国科研数据管理的三个阶段

时间	进展	代表事件
第一阶段（1984～2004）	科研数据的交流和管理阶段	1984 中科院牵头组织“中国 CODATA 委员会”
第二阶段（2004～2014）	科研数据仓储的建设与数据质量控制阶段	2010 科技部“国家科学数据中心网”
第三阶段（2014 以后）	科研数据开放共享的权益管理阶段	2012 全国创新大会上深化科技体制改革的相关要求

表 16 中国图书馆界的科研数据管理服务阶段

时间	进展	具体行动
第一阶段	在数字文献管理的基础上，拓展为科研数据管理	国家图书馆参与文化部“国家文化数字网”项目
第二阶段	协助科研团队实行科研数据管理的仓储建设与咨询服务	CALIS 的“高校科学数据管理机制及管理平台研究”项目
第三阶段	成为科研团队的科研数据开放共享的主要支撑机制	（目前处于第一和第二阶段）

在全球科研数据政策的调研基础上，绘制科研数据权益管理的政策图谱，如图 1 所示。

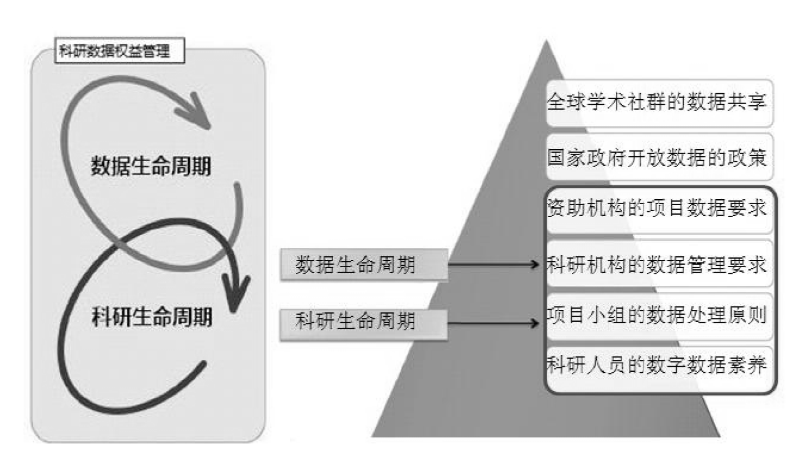


图 1 科研数据权益管理的政策层级

全球学术社群和国家政府所提供的政策文件，主要提供原则性的指导意见；然而，具体实施细节，以资助机构、科研机构、项目小组、科研人员的政策和手册上出现较多。所以，进一步搜索已经实施若干时间、具有一定公信力的机构的综合报告，将这些报告的内容细分为主要内容架构，整理如表 17 所示。

表 17 科研数据权益政策的分析框架

政策文件类型	主要内容架构
《资助机构的项目数据要求分析框架》	机构名称、政策文件、共享状况、强制政策、制裁能力、数据共享成本、数据共享的时间、可获取性
《科研机构的科研数据管理分析框架》	机构名称、学科领域、服务规模、核心成员、建立年代、系统架构、数据共享许可、可获取性等
《项目小组的数据处理原则分析框架》	项目名称、学科领域、行业规范（学术社群、国家、国际）、资助者、项目过程、数据采集原则、数据共享许可、可获取性、附加价值等
《科研人员的科研数据素质分析框架》	机构政策、科研领域的伦理规范、部门政策、开放获取原则、发布科研成果的需求、出版社的要求、与其他项目签订的数据共享协议、国际政策现况、资助机构的规则与要求

所以,为支持图书馆界从事教育科研人员在数据管理上的服务,以及支持图书馆界作为面向创新型社会的知识管理基础设施,科研数据权益管理政策服务的需求日益凸显。为此,需要组织起政策研究咨询团队的服务工作,以期阶段性支持图书馆顺利开展这项新兴业务。

12 未来展望:支持科研生命周期的科研数据权益管理的政策服务

科研数据的生命周期,包括:备份、采集、遴选、典藏、共享等。在每个不同阶段中,均有各个参与者和权益者所关心的不同问题,这些包括:科研人员、学术社群、机构、研究资助者、出版商、公众等的各自相异的权利与利益,对科研数据权益管理提出复杂的要求和挑战,需要未来在实践中逐步梳理和呈现。本文所形成的政策研究框架,对于一般情况下的记录、存储、保存、发布文件的措施,具有参考意义,比如:管理者需要考虑数据重用能力、数据授权许可、开放性、科研数据管理的前提条件等。

综上所述,还有一些普遍性质的议题,在实施科研数据管理以及权益政策制定的过程中,需要考虑。综合考虑的议题,整理如表 18 所示。

表 18 科研数据权益政策的主要元素

数据权益范围[107-108]	数据控制手段[109]	数据开放程度[107,110]
公共知情权利 合法商业利益 国家安全机密 个人隐私权 知情同意权	匿名化 有限访问 封闭化 时滞期 技术保护措施 数字著作权管理	可访问性 可评价性 可解读性 可使用性 可重用性 可操作性

在图 1 中各个政策层级均需重视表 15 的内容,然而它们应该有所不同的重点。全球学术社群和政府监管部门需要充分重视的数据权益范围:

• 公共知情权利(Public right to know):人民有权利知道向政府纳税后的科研调查数据。

• 合法商业利益(Commercial Benefit):从事数据服务的企业有权获得合理利润。

• 国家安全机密(National security secrets):国家出于安全防卫需要,而隐蔽和保护若干数据,通常情况下,只提供结果而不提供推导科研结果过程中的科研数据。

• 个人隐私权(Personal privacy):数据涉及到人类主体(Human bodies)时,需要保护个人隐私,无论是直接披露个人信息或者间接得以为第三方获取个人信息的信息。

• 知情同意权(Informed Consent):数据采集时的对象涉及人类主体时,以及数据访问的记录涉及终端用户时,均需使其有足够完整的信息告知,使其形式自我决策。

资助机构与科研机构需要充分重视的数据控制手段:

• 匿名化(Anonymisation):在数字数据环境中保护隐私的作法。

• 有限访问(limited Access):发布元数据提供检索,但是不提供全文检索或者全文获取,或者通过用户注册机制或者网域识别,对来访用户实施不同程度的限制。

• 封闭化(Closed):在系统上实施开放数据、开放系统架构、开放关联数据、开放平台、知识元库等的接口封闭,使得数据无从释放和重用。

• 时滞期(Embargo):出版发布数据后,在一定时间内,不实施开放共享的时间跨度;不同的出版商、科研机构、资助机构会有不同的时滞期要求,例如 6 个月或者 12 个月。

• 技术保护措施(TPM):通过硬件或软件或者综合两者,以防止或限制复制,以及控制访问和数字内容的技术手段。

• 数字著作权管理(DRM):在格式转换和档案迁移上,透过多重签名认证手段,达到对用户使用内容的限制。

以及项目小组和科研人员需要充分重视的数据开放程度：

・可访问性（accessible）：数据必须配置在能够被发现的实物以及能够被使用的形式上。

・可评价性（accessable）：能够判断数据或者信息的可靠性。数据必需因为不同受众而有差异化。数据必须提供科学工作结果的账目，用以解读和仔细检查它们。

・可解读性（intelligible）：仔细检查某事。受众需要能够对交流内容产生某些判断和评价，他们需要判断论点的本质，需要判断产生这些论点的完整性和可靠性。

・可使用性（useable）：数据或者信息能够被使用的形式。

・可重用性（reuse available）：根据开放许可使得开放数据在使用上没有任何限制。

・可操作性（executable）：采取便捷的、可修改的、以及可以被检索、下载、索引、搜索的开放格式。格式应该是机械可读的（即：数据有合理结构，允许自动化作业处理）。

为了顺利开展科研数据权益政策的服务工作，国家科学图书馆所成立的科技信息政策研究与咨询中心积极面向社会各界，特别是图书馆行业的同仁，寻求共同开发或者纯以学术交流和经验交换为目的的“问题解决方案（Solution）”合作。图 2 显示这种交流合作的循环过程和互助方式。

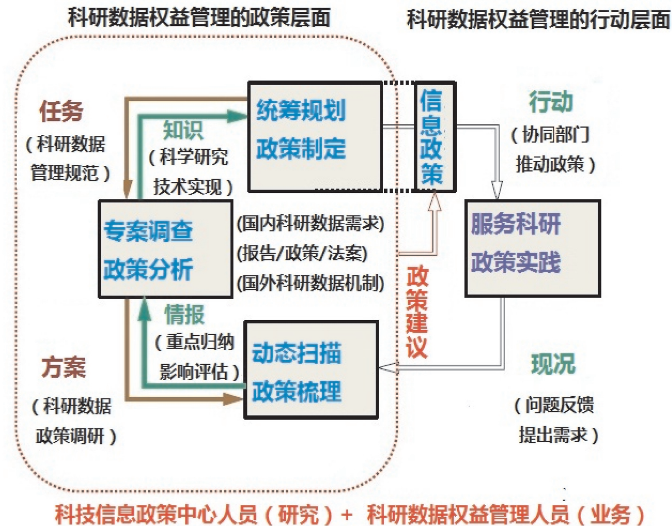


图 2　科研数据权益管理政策的合作模式

科技信息政策研究与服务咨询中心，定期从事国内外相关政策的动态扫描与政策梳理，针对有意义的课题或者重大战略发展的需求，进行专案调查或者政策分析，最终交付具体实施方案于决策管理层，提供统筹规划和政策制定的工作。面向馆内、馆外、院内、院外等不同实施科研数据管理的单位或者部门，提供信息政策模板或者政策建议，同时吸收与学习这些单位的宝贵经验与实施情境。在此过程中，挖掘与发现政策中心在面对实际从事有关工作的合作伙伴时，在经验、理论、方案上的不足之处，持续进行情报、知识、行动的改进，力求能够促使合作伙伴的现况最优化。

根据这套合作模式与图 1 所述的政策层级框架，以及前述的科研数据使用、引用、存储、传播、共享等权益议题等科学数据权益分析的基本框架，未来将会逐一发布经过调研与实践后所形成的各种政策指南。目前，在国家科学图书馆的科技信息政策中心，已经全文翻译了《G8 开放数据宪章与技术附件》[111]、《开放数据行政命令与实施细节》[112-113]、《出版商和大学

对OSTP强制政策的应对措施》[114]、《科研数据的处理原则》[115]、《数据管理计划检查表》[116]、《著作作者的创作共用指南》[117]等一系列政策文件，并且即将发布《图书馆著作权动态》的特刊：科研数据权益管理。未来将有更多的进阶研究。抛砖引玉，是以为文。

致谢：本文初次在"中国图书馆学会专业图书馆分会2013年学术年会"上进行汇报，受到陈锐馆长、刘细文教授、于润升主任等在内的老师们指教；本文受到编辑部的协助。在此致谢。

参考文献

1 The Royal Society. Science as an open enterprise[OL].[2012-06-21]. http://royalsociety.org/policy/projects/science-public-enterprise/report/
2 顾立平译. 英国皇家学会：科学是开放事业(节录)[OL].[2012-06-21]. http://ir.las.ac.cn/handle/12502/6166
3 谢春枝，燕今伟. 国内外高校科学数据管理和机制建设研究[J]. 图书情报工作，2013，57(6)：12-17，38
4 丁宁，马浩琴. 国外高校科学数据生命周期管理模型比较研究及借鉴[J]. 图书情报工作，2013，57(6)：18-22
5 赖剑菲，洪正国. 对高校科学数据管理平台建设的建议[J]. 图书情报工作，2013，57(6)：23-27
6 胡永生，刘颖. 基于用户调查的高校科学数据管理需求分析[J]. 图书情报工作，2013，57(6)：28-32，78
7 刘霞，饶艳. 高校图书馆科学数据管理与服务初探——武汉大学图书馆案例分析[J]. 图书情报工作，2013，57(6)：33-38
8 洪正国，项英. 基于Dspace构建高校科学数据管理平台——以蝎物种与毒素数据库为例[J]. 图书情报工作，2013，57(6)：39-42，84
9 Borgman C.L.. The conundrum of sharing research data[J]. Journal of the American Society for Information Science and Technology，2012，63(6)：1059-1078
10 Borgman C.L.(著)，青秀玲(译). 科研数据共享的挑战[J] 现代图书情报技术，2013(5)：1-20
11 Marchionini G.. 科研数据管理：保障数据质量，促进iSchools新科学研究[J]. 图书情报知识，2013(4)：4-9
12 QinJian. Research Lifecycle，Data Lifecycle，and Data Infrastructure Services[OL].[2013-06-26]. http://jianqin.metadataetc.org/wp-content/uploads/2013/06/LCAS2013_data_service_infrastructure.pdf
13 黄永文，张建勇，黄金霞，等. 国外开放科学数据研究综述[J] 现代图书情报技术，2013(5)：21-27
14 孟祥保，钱鹏. 国外数据管理专业教育实践与研究现状[J]. 中国图书馆学报，2013，39(6)：63-74
15 Garriga-Portola，M.. Open data? Yes，but in a sustainable way [J]. Profesional de la Informacion，2011，20(3)：298-303
16 Hey T.，Tansley S.，Tolle K.. The Fourth Paradigm：Data-Intensive Scientific Discovery[M/OL].[2012-03-08]. http://research.microsoft.com/en-us/collaboration/fourthparadigm/
17 潘教峰，张晓林等译. 第四范式：数据密集型科学发现[M]. 北京：科学出版社，2012：xviii
18 Herrmann GF.. Non-textual information：Contributions by scientific publishers[OL].[2013-03-20]. http://www.nontextualinformation2013.de/images/Downloads/EefkeSmit_GuidoF.Herrmann_lecture.pdf
19 张晓林. 开放获取、开放知识、开放创新推动开放知识服务模式——3O会聚与研究图书馆范式再转变[J]. 现代图书情报技术，2013，29(2)：1-10
20 Carpenter T.. Standards issues related to moving off the page [OL].[2013-03-20]. http://www.nontextualinformation2013.de/images/Downloads/Todd_Carpenter_lecture.pdf
21 Digitale Information. Start[OL].[2013-03-08]. http://www.allianzinitiative.de/de/start/
22 Digitale Information. Forschungsdaten[OL].[2013-03-08]. http://www.allianzinitiative.de/de/start/handlungsfelder/forschungsdaten/
23 Haeussler C.. Information-sharing in academia and the industry：A comparative study[J]. Research Policy2011，40(1)：105-122
24 Pampel H.，Bertelmann R.. Data Policies im Spannungsfeld zwischen Empfehlung und Verpflichtung[C] // Büttner S. et al.(Eds) Handbuch Forschungsdatenmanagement，Bad Honnef：Bock＋Herchen Verlag. 2011：49-61
25 Friesike S.，Bartling S.. Opening Science－The Evolving Guide On How The Internet Is Changing Research，Collaboration and Scholarly Publishing[M]. Heidelberg：Springer，2013：1-500
26 Klump J. et al. Data publication in the open access initiative. Data Science Journal，2006，5：79-83
27 Chavan V.，Penev L.. The data paper：a mechanism to incentivize data publishing in biodiversity science. BMC Bioinformatics，2011，12(S2). doi：10.1186/1471-2105-12-S15-S2
28 Woutersen-Windhouwer S. et al. Enhanced Publications：Linking Publications and Research Data in Digital Repositories[OL].[2013-11-04]. http://dare.uva.nl/aup/nl/record/316849
29 DCC.. Checklist for a Data Management Plan. v.4.0.[EB/OL]. http://www.dcc.ac.uk/resources/data-management-plans
30 Digitale Information. Forschungsdaten[EB/OL].[2013-03-08]. http://www.allianzinitiative.de/de/start/handlungsfelder/forschungsdaten/

31 Johns Hopkins University. Policy on Access and Retention of Research Data and Materials[EB/OL].[2013-03-08]. http://jhuresearch.jhu.edu/Data_Management_Policy.pdf
32 Knight G.. Funder Requirements for Data Management and Sharing[EB/OL].[2013-03-08]. http://researchonline.lshtm.ac.uk/208596/
33 Ruusalepp R.. A Comparative Study of International Approaches to Enabling the Sharing of Research Data[EB/OL].[2013-03-08]. http://www.jisc.ac.uk/publications/reports/2008/nationaldatafinalreport.aspx
34 Knoppers BM. et al. Publishing SNP Genotypes of Human Embryonic Stem Cell Lines: Policy Statement of the International Stem Cell Forum Ethics Working Party[J]. Stem Cell Reviews and Reports, 2011, 7(3):482-484
35 Boyd IL. et al. An International Quiet Ocean Experiment[J]. Ocenography, 2011, 24(2):174-181
36 O'Boyle NM. et al. Open Data, Open Source and Open Standards in chemistry: The Blue Obelisk five years on[J]. Journal of Cheminformatics, 2011, 3:37
37 Gaulton A. et al. ChEMBL: a large-scale bioactivity database for drug discovery[J]. Nucleic ACIDS Research, 2012, 40(D1): D1100-D1107
38 The International Molecular Exchange Consortium. Submit Your Data[OL].[2011-10-27]. http://www.imexconsortium.org
39 Kerrien S. . et al. The IntAct molecular interaction database in 2012[J]. Nucleic ACIDS Research, 2012, 40(D1):D841-D846
40 Gower AC., Spira A., Lenburg ME.. Discovering biological connections between experimental conditions based on common patterns of differential gene expression[J]. BMC Bioinformatics, 2011, 12:381
41 Day N. et al. CrystalEye: automated aggregation, semantification and dissemination of the world's open crystallographic data [J]. Journal of Applied Crystallography, 2012, 45:316-323
42 Fonseca CG. . et al. The Cardiac Atlas Project-an imaging database for computational modeling and statistical atlases of the heart[J]. BIOinformatics, 2011, 27(16):2288-2295
43 Mozilla. Mozilla Public License[OL].[2011-10-22]. http://www.mozilla.org/MPL/MPL-1.1.txt
44 Pettenati MC., Pirri F., Giuli D.. e-Profile Management as a Basic Horizontal Service for the Creation of Specialized e-Services[C]. Morin JH, Ralyté J, Snene M.(eds). Exploring Services Science, Berlin: Springer, 2010, 53:259-263
45 Mukherjee A., Stern S.. Disclosure or secrecy? The dynamics of Open Science[J]. International Journal of Industrial Organization, 2009, 27(3):449-462
46 Ferrer-Sapena A., Peset F., Aleixandre-Benavent, R.. Access to and reuse of public data: open data and open government[J]. Profesional de la Informacion, 2011, 20(3):260-269
47 Lilley B.. The Ordnance Survey OpenData Initiative[J]. Cartographic Journal, 2011, 48(3):179-182
48 McLaren R., Waters R.. Governing Location Information in the UK[J]. Cartographic Journal, 2011, 48(3):172-178
49 Emam K. et al. De-identification Methods for Open Health Data: The Case of the Heritage Health Prize Claims Dataset[J]. Journal of Medical Internet Research, 2012, 14(1):e33
50 Oprea, TI., Taboureau, O., Bologa, CG.. Of possible cheminformatics futures[J]. Journal of Computer-Aided Molecular Design, 2012, 26(1):107-112
51 European Commission. Commission Recommendation on access to and preservation of scientific information[OL].[2013-05－16]. http://ec.europa.eu/research/science-society/document_library/pdf_06/recommendation-access-and-preservation-scientific-information_en.pdf
52 Office of Science and Technology Policy. Increasing Access to the Results of Federally Funded Scientific Research[OL].[2013-05-16]. http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf
53 European Commission. Communication from the Commission to the European Parliament, the European Economic and Social Committee and the Committee of the Regions[OL].[2013-05-16]. http://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf
54 Archambeault J.. Uses and challenges of visualizations for competitive and technological intelligence[OL].[2013-03-20]. http://www.nontextualinformation2013.de/images/Downloads/Jean_Archambeault_lecture.pdf
55 McMahon B.. Bringing crystal structures to life in the scientific literature[OL].[2013-03-20]. http://www.nontextualinformation2013.de/images/Downloads/Brian_McMahon_lecture.pdf
56 Rosemann U., Brase J., Sens I.. Kommentar aus Bibliothekssicht[J]. Zeitschrift für Bibliothekswesen und Bibliographie, 2011, 58:3-4
57 DataCite.startsite[OL].[2011-07-30]. http://data-cite.org
58 Brase J., Blümel I.. Information supply beyond text: non-textual information at the German National Library of Science and Technology (TIB) – challenges and planning[J]. Interlending and Document Supply, 2010, 38(2):108-117
59 Brase J. et al.. 德国国家科技图书馆的发展思考与实践[J]. 图书情报工作, 2013, 57(20):58-61
60 UK. DATA Archive. Create and manage data-research data lifecycle[OL].[2013-03-03]. http://www.data-archive.ac.uk/

create-manage/life-cycle
61 ADMIRAL. Data management plan[OL]. [2013-03-03]. http://imageweb. zoo. ox. ac. uk/wiki/index. php/ADMIRAL _ Data _ Management_Plan_Template
62 UK. DATA Archive. Create and manage data-research data lifecycle[OL]. [2013-03-03]. http://www. data-archive. ac. uk/create-manage/life-cycle
63 National Institute for Health Research. Research Cycle[OL]. [2013-03-03]. http://www. rds-sc. nihr. ac. uk/public-patient-involvement/the-research-process/
64 ICPSR. Guide to Social Science Data Preparation and Archiving [OL]. [2013-03-03]. http://www. icpsr. umich. edu/files/ICPSR/access/dataprep. pdf
65 Simcoe, TS., Graham, SJH., Feldman, MP.. Competing on Standards? Entrepreneurship, Intellectual Property, and Platform Technologies[J]. Journal of Economics and Management Strategy, 2009, 18(3):775-816
66 UK. DATA Archive. Managing and Sharing Data Brochure[OL]. [2013-03-03]. http://www. data-archive. ac. uk/media/2894/managingsharing. pdf
67 顾立平. 开放数据的集成应用研究[J]. 现代图书情报技术, 2012, 228(Z):1-5
68 顾立平. 数字图书馆发展-个性化、开放化、社群化[M]. 北京:科学技术文献出版社, 2013:61-99
69 顾立平. 开放数据计量研究综述[J]. 现代图书情报技术, 2013, 234(6):1-8
70 Gorlitz O., Staab S.. Federated Data Management and Query Optimization for Linked Open Data[J]. New Directions in Web Data Management, 2011, 331:109-137
71 张晓林. 开放获取、开放知识、开放创新推动开放知识服务——3O会聚与研究图书馆范式再转变[J]. 现代图书情报技术, 2013, 230(2):2-10
72 Omitola T. et al. Put in Your Postcode, Out Comes the Data: A Case Study[C]. Aroyo L. et al(Eds). The Semantic Web: Research and Applications, Berlin: Springer, 2010, 6088:318-332
73 Schomburg S.. Publishing Aleph data as Linked Open Data [OL]. [2011-09-22]. http://igelu. org/wp-content/uploads/2011/09/Schomburg_2011_igelu_final. pdf
74 Van de Sompel H., Hochstenbach P.. Reference Linking in a Hybrid Library Environment. Part 3: Generalizing the SFX Solution in the "SFX@Ghent & SFX@LANL" experiment[J/OL]. D-Lib Magazine, 1999, 5(10): http://dx. doi. org/10. 1045/october99-van_de_sompel
75 Van de Sompel H., Hochstenbach P.. Reference Linking in a Hybrid Library Environment. Part 2: SFX, a Generic Linking Solution[J/OL]. D-Lib Magazine, 1999, 5(4): http://dx. doi. org/10. 1045/april99-van_de_sompel-pt2
76 Van de Sompel H., Hochstenbach P.. Reference Linking in a Hybrid Library Environment. Part 1: Frameworks for Linking [J]. D-Lib Magazine, 1999, 5(4): http://dx. doi. org/10. 1045/april99-van_de_sompel-pt1
77 Van de Sompel H., Beit-Arie O.. Open Linking in the Scholarly Information Environment Using OpenURL Framework[J]. D-Lib Magazine, 2001, 7(3): www. dlib. org/dlib/march01/vandesompel/03vandesompel. html
78 Beit-Arie O.. Linking to the Appropriate Copy: Report of a DOI-Based Prototype[J]. [2011-05-30]. D-Lib Magazine, 2001, 7(9): http://www. dlib. org/dlib/september01/caplan/09caplan. html
79 Hofmann-Apitius M.. Innovation beyond text – the future of scientific communication[OL]. [2013-03-20]. http://www. nontextualinformation2013. de/images/Downloads/Martin _ Hofmann-Apitius_keynote. pdf
80 Altmetric. Fetching details by Altmetric ID[OL]. [2013-04-07]. http://api. altmetric. com/docs/call_id. html
81 Altmetric. Fetching details by DOI[OL]. [2013-04-07]. http://api. altmetric. com/docs/call_doi. html
82 Altmetric. Fetching details by PubMed ID[OL]. [2013-04-07]. http://api. altmetric. com/docs/call_pmid. html
83 Altmetric. Fetching details by arXiv ID[OL]. [2013-04-07]. http://api. altmetric. com/docs/call_arxiv. html
84 Altmetric. Fetching details by ADS bibcode[OL]. [2013-04-07]. http://api. altmetric. com/docs/call_ads. html
85 Chandler A.. NISO IOTA: Improving OpenURLs Through Analytics, in Context[OL]. [2012-09-22]. http://ecommons. cornell. edu/bitstream/1813/19495/2/ATG%20NISO%20IOTA%20in%20context. pdf
86 Culling J.. Link Resolvers and the Serials Supply Chain[OL]. [2007-12-14]. http://www. uksg. org/projects/linkfinal
87 Jewell T. et al. Making Good on the Promise of ERM: A Standards and Best Practices Discussion Paper[OL]. [2012-03-08]. http://www. niso. org/apps/group _ public/download. php/7946/Making_Good_on_the_Promise_of_ERM. pdf
88 Simcoe TS., Graham SJH., Feldman MP.. Competing on Standards? Entrepreneurship, Intellectual Property, and Platform Technologies[J]. Journal of Economics and Management Strategy, 2009, 18(3):775-816
89 Smartcopying. Technological Protection Measures and the Copyright Amendment Act 2006[OL]. [2013-03-08]. http://www. smartcopying. edu. au/copyright-guidelines/hot-topics/technological-protection-measures
90 Kerr IR.. Technological Protection Measures: Part II — The Le-

gal Protection of TPMs[OL].[2013-03-08]. http://papers.ssrn.com/sol3/papers.cfm? abstract_id=705081
91 Alspaugh TA., Asuncion HU., Scacchi W.. Analyzing Software Licenses in Open Architecture Software Systems[OL].[2013-11-04].http://dl.acm.org/citation.cfm? id=1572203
92 Yun S. et al. The Biometric Based Convertible Undeniable Multi-Signature Scheme to Ensure Multi-Author Copyrights and Profits [J]. Wireless Personal Communications, 2011, 60(3):405-418
93 Eschenfelder KR., Desai AC., Downey G.. The Pre-Internet Downloading Controversy: The Evolution of Use Rights for Digital Intellectual and Cultural Works[J]. Information Society, 2011, 27(2):69-91
94 Pickett C.. Eliminating Administrative Churn: The "Big Deal" and Database Subscriptions[J]. Serials Review, 2011, 37(4): 258-261
95 Corbett S.. Creative Commons Licences, the Copyright Regime and the Online Community: Is there a Fatal Disconnect? [J]. Modern Law Review, 2011, 74(4):503-531
96 Torres-Padrosa V., Delgado-Merce J.. Alternatives for the self-management of author rights in the digital world[J]. Profesional de la Informacion, 2011, 20(1):61-69
97 Torres-Salinas D., Robinson-Garcia N., Cabezas-Clavijo A.. Sharing scientific research data: introduction to data sharing [J]. Profesional de la Informacion, 2012, 21(2):173-184
98 Parsons MA. et al. A conceptual framework for managing very diverse data for complex, interdisciplinary science[J]. Journal of Information Science, 2011, 37(6):555-569
99 Coppens S. et al. Toward interoperable bioscience data[J]. Nature Genetics, 2012, 44(2):121-126
100 Reichman OJ., Jones MB., Schildhauer MP.. Challenges and Opportunities of Open Data in Ecology[J]. Science, 2011, 331 (6018):703-705
101 Goovaerts M.. Oceandocs and Open Science Dirctory: Two Facets of the Information Policy of UNESCO/IOC-IODE[C]. Confluence of Ideas: Evolving to Meet the Challenges of Global Change, 2010, 97:109
102 Ramachandran R. et al. Talkoot: software tool to create collaboratories for earth science[J]. Earth Science Informatics, 2012, 5(1):33-41
103 Jonjic T.. Juggling Between Open Science and the Market: Public Science Responses to the Patentability of Biomedical Research Tools [J]. Periodicum Biologorum, 2010, 112(4):381-390
104 Penin J.. On the Consequences of Patenting University Research: Lessons from a Survey of French Academic Inventors [J]. Industry and Innovation, 2010, 17(5):445-468
105 Grubb AM., Easterbrook, SM.. On the Lack of Consensus over the Meaning of Openness: An Empirical Study[J]. PLOS ONE, 2011, (8):e23420.doi:10.1371/journal.pone.0023420
106 Kindling M., Schirmbacher P.. Die digitale Forschungswelt als Gegenstand der Forschung. Wissenschaft und Praxis, 2013, 64 (2/3):127-136
107 Die Allianz der deutschen Wissenschaftsorganisationen. Grundsätze zum Umgang mit Forschungsdaten[EB/OL].[2010-06-24]. http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/grundsaetze/
108 The White House. National Strategy for Trusted Identities in Cyberspace[EB/OL].[2011-04-07].http://www.whitehouse.gov/sites/default/tiles/rss viewer/NSTICstrategv 04 15Il .pdf
109 NIST. Standards for Security Categorization of Federal Information and Information Systems[EB/OL].http://csrc.nist.gov/publications/fips/fips199/FfPS-PUB-199-final.pdf/
110 President Barack Obama.Memorandum on Building a 21st Century Digital Government[EB/OL].[2012-05-23]. http://www.whitehouse.gov/sites/default/liles/uploads/2012digitalmemrei.pdf/
111 Cabinet Office. G8 Open Data Charter and Technical Annex [EB/OL].[2013-06-19]. https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex
112 White House. Executive Order — Making Open and Machine Readable the New Default for Government Information[EB/OL]. [2013-05-04]. http://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-
113 White House.Open Data Policy-Managing Information as an Asset[EB/OL].[2013-05-04]. http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf
114 Covey T.D.. Publishers and universities respond to the OSTP mandate[EB/OL].[2013-07-10]. http://works.bepress.com/denise_troll_covey/72
115 Die Allianz der deutschen Wissenschaftsorganisationen. Grundsätze zum Umgang mit Forschungsdaten[EB/OL].[2010-06-24]. http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/grundsaetze/
116 DCC.. Checklist for a Data Management Plan. v.4.0.[EB/OL]. http://www.dcc.ac.uk/resources/data-management-plans
117 Collins E., Milloy C., Stone G. et al. Guide to Creative Commons[EB/OL].[2013-07-15]. http://oapen-uk.jiscebooks.org/ccguide/

(收稿日期:2013-09-06)